

OY 32 DROEYECROCMOLETSGOMR-CVSQCDKRFEDIDWSK 71

RESULT 2

ID VCLA.GOSHI STANDARD; PRT; 605 AA.

AC P09799;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE VICILIN GC72-A PRECURSOR (ALPHA-GLOBULIN A).

OS Gossypium hirsutum (Upland cotton).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

OC euhylophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

RN [1]

RP SEQUENCE FROM N.A.

RA Chlan C.A., Borroto K., Kamalay J.A., Dure L. III;

RT "Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Sequences and genomic organization of the alpha globulin (VICILIN) genes of cottonseed.";

RL Plant Mol. Biol. 9:533-546(1987).

CC -1- FUNCTION: SEED STORAGE PROTEIN.

CC -1- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN BODIES.

CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASCOLIN, VICILIN, CONVICILIN, CONGLICININ, ETC.).

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M19378; AAA33069.1; -

DR PIR: S06398; S06398.

DR HSSP: P50477; ICAX.

DR PRAM: PF00546; Seedstore_7s; 1.

KW Seed storage protein; Signal.

FT SIGNAL 1 23

FT CHAIN 24 605 VICILIN GC72-A.

SO SEQUENCE 605 AA; 71049 MW; C9DB9371C976953B CRC64;

Query Match 28.0%; Score 101; DB 1; Length 605;

Best Local Similarity 39.0%; Pred. No. 1.36e-05;

Matches 16; Conservative 12; Mismatches 9; Indels 4; Gaps 4;

DB 33 SEDDPOORVEDCRRC-OLETRGOTEDKCDRESEQLAKE 72

OY 29 SEDRDE-YECCRQCMOLETSGOM-R-RCVSQCDKRFED 66

RESULT 3

ID AGRP.LUCFY STANDARD; PRT; 47 AA.

AC P56568;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE 6.5 KD ARGININE/GLUTAMATE-RICH POLYPEPTIDE (6.5K-AGRP).

OS Lufla cylindrica (Smooth loofah) (Sponge gourd).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Cucurbitales; Cucurbitaceae; Lufla.

OC [1]

RN SEQUENCE

RP TISSUE-SEED.

RX MEDLINE: 97357433.

RA Kimura M., Park S.S., Sakai R., Yamasaki N., Funatsu G.;

RT "Primary structure of 6.5K-arginine/glutamate-rich polypeptide from the seeds of sponge gourd (Lufla cylindrica).";

RL Biosci. Biotechnol. Biochem. 61:984-988(1997).

CC -1- FUNCTION: STORAGE PROTEIN WHICH PROVIDES NITROGEN AND CARBON RESERVES DURING GERMINATION AND SEEDLING GROWTH.

CC -1- MASS SPECTROMETRY: MW=5693.39; METHOD=MALDI.

CC -1- SIMILARITY: SOME, TO 7S SEED STORAGE PROTEINS.

KW Seed storage protein.

FT DISULFID 12 33

FT DISULFID 16 29

SO SEQUENCE 47 AA; 5698 MW; 588B0EC82273AC05 CRC64;

Query Match 27.7%; Score 100; DB 1; Length 47;

Best Local Similarity 41.2%; Pred. No. 2.12e-05;

Matches 14; Conservative 6; Mismatches 13; Indels 1; Gaps 1;

DB 6 RTEYACRVACVAEHGVERORRCQVCEKRLRE 39

OY 33 ROEYECRCROCMOLETSGOM-RCVSCDCKRFE 65

RESULT 4

ID TSPI.HUMAN STANDARD; PRT; 1170 AA.

AC P07996;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE THROMBOSPONDIN 1 PRECURSOR.

GN THBS1 OR TSPI OR TSP.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-ENDOTHELIAL CELLS.

RX MEDLINE: 87057617.

RA Lawler J., Hynes R.O.;

RT "The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologs with several different proteins.";

RT J. Cell Biol. 103:1635-1648(1986).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE: 89133590.

RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L., Baungartel D.M., Rotwein P., Frazier W.A.;

RT "Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.";

RT J. Cell Biol. 108:729-736(1989).

RN [3]

RP SEQUENCE OF 1-397 FROM N.A.

RX MEDLINE: 87157592.

RA Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;

RT "Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.";

RT Biochemistry 25:8418-8425(1986).

RN [4]

RP SEQUENCE OF 1-374 FROM N.A.

RX MEDLINE: 86287276.

RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;

RT "Characterization of a cDNA encoding the heparin and collagen binding domains of human thrombospondin.";

RT Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).

RN [5]

RP SEQUENCE OF 1-166 FROM N.A.

RX MEDLINE: 89291870.

RA Laherty C.D., Gierman T.M., Dixit V.M.;

RT "Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.";

RT J. Biol. Chem. 264:11222-11227(1989).

RN [6]

RP SEQUENCE OF 1028-1170 FROM N.A.

RA la Fleur M., Jbhin C., Gauchier J., Kreis C.G.;

RT Submitted (XXX-1992) to the EMBL/Genbank/DBJ databases.

RL -1- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND

[illegible]

```

FT CAROBYD 1067 1067 POTENTIAL. T -> A (IN REF. 2, 3 AND 4).
FT CONFLICT 84 84 T -> A (IN REF. 2).
FT CONFLICT 523 523
SQ SEQUENCE 1170 AA, 129412 MW, 69B3E5DE5AE3A395E CRC64;

Query Match 26.3%: Score 95; DB 1; Length 1170;
Best Local Similarity 50.0%; Pred. 1,90e-04;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Db 416 SSVQRTCHIDCECKRFRKDGWS 439
      :| | | :| | | :| |
QY 48 TSGQMRC-VSQCKRFEEDIMWS 70

RESULT 5
ID TSP1_MOUSE 1 STANDARD: PRT; 1170 AA.
AC P35441.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE THROMBOSPONDIN 1 PRECURSOR.
GN THS1 OR TSP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE: 92128941.
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.:
RL "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE: 92147683.
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.:
RL "Characterization of mouse thrombospondin 2 sequence and expression
RL during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RN SEQUENCE OF 1-490 FROM N.A.
RX MEDLINE: 90375546.
RA Bornstein P., Afifi D., Devarayalu S., Framson P., Li P.:
RL "Characterization of the mouse thrombospondin gene and evaluation of
RL the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16696(1990).
CC -1- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND
CC CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,
CC LAMININ AND TYPE V COLLAGEN.
CC -1- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.
CC -1- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WFC DOMAIN.
CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS
CC (WHICH BIND CALCIUM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use: by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62470; AAA50611.1; JOINED.
DR EMBL; M62450; AAA50611.1; JOINED.
DR EMBL; M62451; AAA50611.1; JOINED.
DR EMBL; M62452; AAA50611.1; JOINED.
DR EMBL; M62453; AAA50611.1; JOINED.
DR EMBL; M62454; AAA50611.1; JOINED.
DR EMBL; M62455; AAA50611.1; JOINED.
DR EMBL; M62456; AAA50611.1; JOINED.

```

DR	EMBL	M62457	AAA50611.1	JOINED.	
DR	EMBL	M62458	AAA50611.1	JOINED.	
DR	EMBL	M62459	AAA50611.1	JOINED.	
DR	EMBL	M62460	AAA50611.1	JOINED.	
DR	EMBL	M62461	AAA50611.1	JOINED.	
DR	EMBL	M62462	AAA50611.1	JOINED.	
DR	EMBL	M62463	AAA50611.1	JOINED.	
DR	EMBL	M62464	AAA50611.1	JOINED.	
DR	EMBL	M62465	AAA50611.1	JOINED.	
DR	EMBL	M62466	AAA50611.1	JOINED.	
DR	EMBL	M62467	AAA50611.1	JOINED.	
DR	EMBL	M62468	AAA50611.1	JOINED.	
DR	EMBL	M62469	AAA50611.1	JOINED.	
DR	EMBL	M87276	AAA53063.1	-	
DR	EMBL	J05606	AAA40431.1	-	
DR	EMBL	J05605	AAA40431.1	JOINED.	
DR	PIR	A40558			
DR	PIR	B42587			
DR	PIR	A37905			
DR	HSSP	P35555	LEMO.		
DR	DR	MG1	98737	THBS1.	
DR	PROSITE	PS00022	EGF_1	FALSE_NEG.	
DR	PROSITE	PS01186	EGF_2	1.	
DR	PROSITE	PS01208	WVEC	1.	
DR	PFAM	PF00008	EGF	2.	
DR	PFAM	PF00090	tsp_1	3.	
DR	PFAM	PF00093	WVC	1.	
KW	Glycoprotein				
KW	EGF-like domain				
FT	SIGNAL	1	18	POTENTIAL.	
FT	CHAIN	19	1170	THROMBOSPONDIN 1.	
FT	DOMAIN	19	232	HEPARIN-BINDING (POTENTIAL).	
FT	DOMAIN	316	373	WVEC.	
FT	DOMAIN	379	548	3 X TSP TYPE-1 REPEATS (CS-LIKE).	
FT	DOMAIN	549	690	3 X EGF-TYPE REPEATS.	
FT	DOMAIN	723	950	7 X TSP TYPE-3 REPEATS (CA-BINDING).	
FT	DOMAIN	951	1170	C-TERMINAL.	
FT	REPEAT	379	430	TSP TYPE-1 1.	
FT	REPEAT	435	491	TSP TYPE-1 2.	
FT	REPEAT	492	548	TSP TYPE-1 3.	
FT	DOMAIN	549	587	EGF-LIKE 1.	
FT	DOMAIN	588	645	EGF-LIKE 2.	
FT	DOMAIN	646	690	CALCIUM-BINDING (POTENTIAL).	
FT	REPEAT	723	758	EGF-LIKE 3.	
FT	REPEAT	759	781	TSP TYPE-3 1.	
FT	REPEAT	782	817	TSP TYPE-3 2.	
FT	REPEAT	818	840	TSP TYPE-3 3.	
FT	REPEAT	841	878	TSP TYPE-3 4.	
FT	REPEAT	879	914	TSP TYPE-3 5.	
FT	REPEAT	915	950	TSP TYPE-3 6.	
FT	REPEAT	951	950	TSP TYPE-3 7.	
FT	SITE	926	928	CELL ATTACHMENT SITE (POTENTIAL).	
FT	DISULFID	270	270	INTERCHAIN (PROBABLE).	
FT	DISULFID	274	274	INTERCHAIN (PROBABLE).	
FT	DISULFID	551	562	BY SIMILARITY.	
FT	DISULFID	556	572	BY SIMILARITY.	
FT	DISULFID	575	586	BY SIMILARITY.	
FT	DISULFID	592	608	BY SIMILARITY.	
FT	DISULFID	599	617	BY SIMILARITY.	
FT	DISULFID	620	644	BY SIMILARITY.	
FT	DISULFID	650	663	BY SIMILARITY.	
FT	DISULFID	657	676	BY SIMILARITY.	
FT	DISULFID	678	689	BY SIMILARITY.	
FT	CARBOHYD	248	248	POTENTIAL.	
FT	CARBOHYD	360	360	POTENTIAL.	
FT	CARBOHYD	708	708	POTENTIAL.	
FT	CARBOHYD	1067	1067	POTENTIAL.	
FT	CONFUCT	1025	1025	F -> L (IN AAA53063).	
FT	SEQUENCE	1170	129646	MM; 0435E93615E7F06	
Q1				CRC64;	

DB	4416	SV0TRTCHIOECCRRKFGDGM	439
QY	48	TSGCMRRC-VSCCCKRFEEDIDMS	70
RESULT	6	STANDARD:	PRT: 1173 AA.
ID	TSPL_XENMLA		
AC	P35448;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DT	01-OCT-1996 (Rel. 34, Last annotation update)		
DE	THROMBOSPONDIN 1 PRECURSOR.		
GN	THBS1 OR TSPL.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;		
OC	Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;		
OC	Xenopus.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;		
RL	Submitted (xxx-1993) to the EMBL/GenBank/DBJ databases.		
CC	- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND		
CC	CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,		
CC	LAMININ AND TYPE V COLLAGEN.		
CC	- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.		
CC	- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.		
CC	- SIMILARITY: CONTAINS 1 WVC DOMAIN.		
CC	- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.		
CC	- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS		
CC	(WHICH BIND CALCIUM).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).		
CC	-----		
CC	EMBL; L04278; -; NOT_ANNOTATED_CDS.		
DR	HSSP; P00740; IIXA.		
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.		
DR	PROSITE; PS01186; EGF_2; 1.		
DR	PROSITE; PS01208; WVC; 1.		
DR	Pfam; PF00008; EGF; 2.		
DR	Pfam; PF00090; tsp_1; 3.		
DR	Pfam; PF00093; wvc; 1.		
KW	Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;		
KW	EGF-like domain; Signal.		
FT	1	22	POTENTIAL.
FT	CHAIN		THROMBOSPONDIN 1.
FT	DOMAIN	23	HEPARIN-BINDING (POTENTIAL).
FT	DOMAIN	23	WVC.
FT	DOMAIN	319	3 X TSP TYPE-1 REPEATS (CS-LIKE).
FT	DOMAIN	382	3 X EGF-TYPE REPEATS.
FT	DOMAIN	550	3 X TSP TYPE-3 REPEATS (CA-BINDING).
FT	DOMAIN	726	C-TERMINAL.
FT	DOMAIN	954	TSP TYPE-1 1.
FT	REPEAT	382	TSP TYPE-1 2.
FT	REPEAT	438	TSP TYPE-1 3.
FT	REPEAT	438	EGF-LIKE 1.
FT	REPEAT	494	EGF-LIKE 2.
FT	REPEAT	495	EGF-LIKE 3.
FT	DOMAIN	550	EGF-LIKE 3.
FT	DOMAIN	591	EGF-LIKE 2.
FT	DOMAIN	648	EGF-LIKE 2.
FT	REPEAT	726	TSP TYPE-3 1.
FT	REPEAT	762	TSP TYPE-3 2.
FT	REPEAT	762	TSP TYPE-3 3.
FT	REPEAT	784	TSP TYPE-3 4.
FT	REPEAT	820	TSP TYPE-3 5.
FT	REPEAT	821	TSP TYPE-3 6.
FT	REPEAT	844	TSP TYPE-3 7.
FT	REPEAT	882	TSP TYPE-3 7.
FT	REPEAT	918	TSP TYPE-3 7.
FT	SITE	929	CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID	931	BY SIMILARITY.
FT	DISULFID	554	BY SIMILARITY.
FT	DISULFID	565	BY SIMILARITY.
FT	DISULFID	559	BY SIMILARITY.
FT	DISULFID	575	BY SIMILARITY.

FT DISULFID 578 589 BY SIMILARITY.
 FT DISULFID 595 611 BY SIMILARITY.
 FT DISULFID 602 620 BY SIMILARITY.
 FT DISULFID 623 647 BY SIMILARITY.
 FT DISULFID 653 666 BY SIMILARITY.
 FT DISULFID 660 679 BY SIMILARITY.
 FT DISULFID 681 692 BY SIMILARITY.
 FT CARBOHYD 155 155 POTENTIAL.
 FT CARBOHYD 158 158 POTENTIAL.
 FT CARBOHYD 250 250 POTENTIAL.
 FT CARBOHYD 363 363 POTENTIAL.
 FT CARBOHYD 705 705 POTENTIAL.
 FT CARBOHYD 711 711 POTENTIAL.
 FT CARBOHYD 1070 1070 POTENTIAL.
 SQ SEQUENCE 1173 AA; 130019 MW; A9F036D6516C0F24 CRC64;

Query Match 26.0%; Score 94; DB 1; Length 1173;
 Best Local Similarity 50.0%; Pred. No. 2,94e-04;
 Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Db 419 SSVQTRSCIOCDKRRFKDGGWS 442
 48 TSGQMRRC-VSGCCKRFEEDIDWS 70

RESULT 7
 ID SBF-SOYBN STANDARD; PRT; 524 AA.
 AC 004672;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE SUCROSE-BINDING PROTEIN PRECURSOR (SBP).
 GN SBP.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC eumphylophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
 OC Glycine.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-50.
 RC TISSUE-EMBRYO;
 RX MEDLINE; 93104680.
 RA Grimes H.D., Overvoorde P.J., Ripp K., Franceschi V.R., Hitz W.D.;
 RT "A 62-kD sucrose binding protein is expressed and localized in
 tissues actively engaged in sucrose transport.";
 RL Plant Cell 4:1561-1574(1992).
 CC -1- FUNCTION: PLAYS A ROLE IN SUCROSE TRANSPORT.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.
 CC SEVERAL CELL TYPES ENGAGED IN THE PLASMA MEMBRANE OF
 CC MESOPHYLL CELLS OF YOUNG SINK LEAVES, THE COMPANION CELLS OF
 CC MATURE PHLOEM AND THE CELLS OF DEVELOPING COTYLEDONS.
 CC -1- DEVELOPMENTAL STAGE: IN THE COTYLEDON, EXPRESSION IS NOT DETECTED
 CC UNTIL 10 DAYS AFTER FERTILIZATION. BETWEEN 10-19 DAYS AFTER
 CC FERTILIZATION, EXPRESSION INCREASES RAPIDLY BUT DECLINES 20-30
 CC DAYS AFTER FERTILIZATION. 30 DAYS AFTER FERTILIZATION, NO
 CC EXPRESSION OCCURS. THIS EXPRESSION PATTERN CLOSELY PARALLELS THE
 CC RATE OF SUCROSE UPTAKE IN THE COTYLEDON.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L06038; AAB03894.1; -;
 DR EMBL; J01730; J01730.
 DR PIR; P50477; ICAM.
 DR HSSP; P50477; ICAM.
 DR PFWM; PFW00546; Seedstore_7s; 1.
 DR Transport; Sugar transport; Signal; Membrane.
 FT SIGNAL 1 29

FT CHAIN 30 524 SUCROSE-BINDING PROTEIN.
 SQ SEQUENCE 524 AA; 60522 MW; 0251EE90796EF341 CRC64;

Query Match 21.3%; Score 77; DB 1; Length 524;
 Best Local Similarity 44.4%; Pred. No. 3.08e-01;
 Matches 12; Conservative 6; Mismatches 6; Indels 3; Gaps 2;

Db 43 CKHOCQOQOYTEGKRVCLOSCD-RY 68
 39 CKRCQMLE-TSGQMRRCVSGCCKRF 63

RESULT 8
 ID TSP2_HUMAN STANDARD; PRT; 1172 AA.
 AC P35442;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-1999 (Rel. 38, Last annotation update)
 DE THROMBOSPONDIN 2 PRECURSOR.
 GN THBS2 OR TSP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 94010892.
 RA Labell T.L., Byers P.H.;
 RT "Sequence and characterization of the complete human thrombospondin 2
 cDNA: potential regulatory role for the 3' untranslated region.";
 RL Genomics 17:223-228(1993).
 RN [2]
 RP SEQUENCE OF 560-1172 FROM N.A.
 RC TISSUE-FIBROBLAST;
 RX MEDLINE; 92217961.
 RA Labell T.L., McGockey Milewicz D.J., Distche C.M., Byers P.H.;
 RT "Thrombospondin II: partial cDNA sequence, chromosome location, and
 RT expression of a second member of the thrombospondin gene family in
 RT humans.";
 RL Genomics 12:421-429(1992).
 CC -1- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND
 CC CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,
 CC LAMININ AND TYPE V COLLAGEN.
 CC -1- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WFEC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS
 CC (WHICH BIND CALCIUM).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L12350; AAA03703.1; -;
 DR EMBL; M61339; -; NOT_ANNOTATED_CDS.
 DR PIR; A42173; A42173.
 DR HSSP; P00740; IIXA.
 DR MIM; 188061; -;
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01208; WFEC; 1.
 DR PFWM; PFW0008; EGF; 2.
 DR PFWM; PFW0090; tsp_1; 3.
 DR PFWM; PFW0093; vwc; 1.
 DR Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 EGF-like domain; signal.
 FT SIGNAL 1 18
 FT CHAIN 19 1172
 FT DOMAIN 19 232 THROMBOSPONDIN 2.
 HEPARIN-BINDING (POTENTIAL).

FT DOMAIN 318 375 WMFC.
FT DOMAIN 381 548 3 X TSP TYPE-1 REPEATS (CS-LIKE).
FT DOMAIN 549 692 3 X EGF-TYPE REPEATS.
FT DOMAIN 725 952 7 X TSP TYPE-3 REPEATS (CA-BINDING).
FT DOMAIN 953 1172 C-TERMINAL.
FT REPEAT 381 432 TSP TYPE-1.1.
FT REPEAT 437 493 TSP TYPE-1.2.
FT REPEAT 494 548 TSP TYPE-1.3.
FT DOMAIN 549 589 EGF-LIKE 1.
FT DOMAIN 590 647 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 648 692 EGF-LIKE 3.
FT REPEAT 725 760 TSP TYPE-3.1.
FT REPEAT 761 783 TSP TYPE-3.2.
FT REPEAT 784 819 TSP TYPE-3.3.
FT REPEAT 820 842 TSP TYPE-3.4.
FT REPEAT 843 880 TSP TYPE-3.5.
FT REPEAT 881 916 TSP TYPE-3.6.
FT REPEAT 917 952 TSP TYPE-3.7.
FT SITE 928 930 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 266 266 INTERCHAIN (PROBABLE).
FT DISULFID 270 270 INTERCHAIN (PROBABLE).
FT DISULFID 553 564 BY SIMILARITY.
FT DISULFID 558 574 BY SIMILARITY.
FT DISULFID 577 588 BY SIMILARITY.
FT DISULFID 594 610 BY SIMILARITY.
FT DISULFID 601 619 BY SIMILARITY.
FT DISULFID 622 646 BY SIMILARITY.
FT DISULFID 652 665 BY SIMILARITY.
FT DISULFID 659 678 BY SIMILARITY.
FT DISULFID 680 691 BY SIMILARITY.
FT CARBOHYD 151 151 POTENTIAL.
FT CARBOHYD 316 316 POTENTIAL.
FT CARBOHYD 330 330 POTENTIAL.
FT CARBOHYD 457 457 POTENTIAL.
FT CARBOHYD 584 584 POTENTIAL.
FT CARBOHYD 710 710 POTENTIAL.
FT CARBOHYD 1069 1069 POTENTIAL.
SQ SEQUENCE 1172 AA: 129955 MW: 2AC7BB230E44C6F5 CRC64;

Query Match 21.3%; Score 77; DB 1; Length 1172;
Best Local Similarity 46.2%; Pred. No. 3.08e-01;

Matches 12; Conservative 3; Mismatches 10; Indels 1; Gaps 1;

Db 416 LGPSIOTRACSLSKDTRINODGWS 441

QY 46 LETSGQWRC-VSCDCKRFEEDIDMS 70

RESULT 9
ID YAZ1_SCHPO STANDARD: PRT: 431 AA.
AC Q10223;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 49.6 KD PROTEIN C13D6.01 IN CHROMOSOME 1.
GN SPAC13D6.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetaceae; Schizosaccharomyces.
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RA Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: 269725; CAA93541.1; -
KW Hypothetical protein.
SQ SEQUENCE 431 AA: 49623 MW: 132FB0915646AC9F CRC64;

Query Match 20.8%; Score 75; DB 1; Length 431;
Best Local Similarity 30.0%; Pred. No. 6.59e-01;

Matches 12; Conservative 11; Mismatches 13; Indels 4; Gaps 4;

Db 46 FKPTFDEEV-HCLA-RVRA-MKRLVSGSN-FDEEDNN 81

QY 31 FDRQVEECKRCQMOLETSQMRRCVSCDCKRFEEDIDMS 70

RESULT 10
ID CYSK_YEAST STANDARD: PRT: 393 AA.
AC P53206;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PUTATIVE CYSTEINE SYNTHASE (EC 4.2.99.8) (O-ACETYL-SERINE
DE SULFHYDRYLASE) (O-ACETYL-SERINE (THIO)-LYASE) (CSASE).
GN YGR012W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE: 97435481.
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII.*;
RL Yeast 13:1077-1090(1997).
CC -I- CATALYTIC ACTIVITY: O-ACETYL-L-SERINE + H(2)S = L-CYSTEINE +
CC ACETATE.
CC -I- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -I- PATHWAY: CYSTEINE BIOSYNTHESIS.
CC -I- SIMILARITY: BELONGS TO THE CYSTEINE SYNTHASE/CYSTATHIONINE BETA-
CC SYNTHASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: 272797; CAA96995.1; -
DR PROSITE: PS00901; CYS_SYNTHASE; 1.
DR PFM: PFO0291; S_T_dehydratase; 1.
KW Hypothetical protein; Lyase; Cysteine biosynthesis;
KW Pyridoxal phosphate.
FT BINDING 86 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 393 AA: 42801 MW: 76C73396D7B69AC CRC64;

Query Match 20.5%; Score 74; DB 1; Length 393;
Best Local Similarity 33.3%; Pred. No. 9.59e-01;
Matches 13; Conservative 8; Mismatches 15; Indels 3; Gaps 3;

Db: 167 DPNQYNAAKKACNELKSGNIGRAVFA-D-OFEEDANN 203

QY 32 DROEV-EECKRCQMOLETSQMRRCVSCDCKRFEEDIDM 69

RESULT 11
ID TSP2_BOVIN STANDARD: PRT: 1170 AA.
AC Q95116;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE THROMBOSPONDIN 2 PRECURSOR (CORTICOTROPIN-INDUCED SECRETED PROTEIN)
DE (CISP).

